

СПИСОК ЛИТЕРАТУРЫ:

1. Ткаченко Т.Я. Инструментальная среда системотехнического обслуживания сложных объектов. – Екатеринбург: ГОУ ВПО «УГТУ-УПИ», 2002. – 203с.
2. Гольдштейн С.Л., Кудрявцев А.Г. Разрешение проблемных ситуаций при поддержке систем, основанных на знаниях: Учеб. пособие. – Екатеринбург: ИД «ПироговЪ», 2006. – 218 с.
3. Гольдштейн С.Л., Кудрявцев А.Г. Структура и технологии системного интеллектуального подсказчика по разрешению проблемных ситуаций // Наука и производство: Сборник научных трудов. – Челябинск: ЧНЦ РАН, 2007. – С. 236 -255.
4. Гольдштейн С.Л., Кудрявцев А.Г. Система наполнения и обнаружения знаний для системного интеллектуального подсказчика. (Статья в этом сборнике)
5. Овдей О.М., Проскудина Г.Ю. Обзор инструментов инженерии онтологий; www.rcdl.ru/papers/2005/sek3_2_paper.pdf
6. Гаврилова Т.А., Хорошевский В.Ф. Базы знаний интеллектуальных систем: Учеб. пособие для вузов. - СПб.; М.; Харьков; Минск: ПИТЕР, 2000. - 384 с.

Гольдштейн С.Л., Кудрявцев А.Г.

СИСТЕМА НАПОЛНЕНИЯ И ОБНАРУЖЕНИЯ ЗНАНИЙ ДЛЯ СИСТЕМНОГО ИНТЕЛЛЕКТУАЛЬНОГО ПОДСКАЗЧИКА

vtsl@dpt.ustu.ru

ГОУ ВПО УГТУ-УПИ

г. Екатеринбург

Разработана исследовательская версия системы наполнения и обнаружения знаний (как важной составной части перспективного системного интеллектуального подсказчика по разрешению проблемных ситуаций со сложными объектами).

The exploratory version of knowledges filling and finding system (as important component part of perspective system intellectual tutor on the permit of problem situations solving with complex objects) is designed.

В рамках реализации проекта по созданию советующей системы нового типа – системного интеллектуального подсказчика (СИП) [1-3] разработана исследовательская версия его важной составляющей, именно, онтолингвистической системы [4] наполнения и обнаружения знаний (СНОЗ СИП), способной как создавать систему (базу) знаний (СЗ) для СИП, так и принимать естественно-языковые запросы с выдачей прямого и развернутого текстовых ответов.

Решение по структуре СНОЗ СИП показано на рис. 1 и 2.

Функции СНОЗ СИП:

- обеспечивать создание СЗ на основе входных текстов;

- принимать естественно-языковые запросы до двух предложений;
- обеспечивать ответы на запросы;
- знакомить обучаемых с технологиями создания СЗ и лингвистического обнаружения знаний.

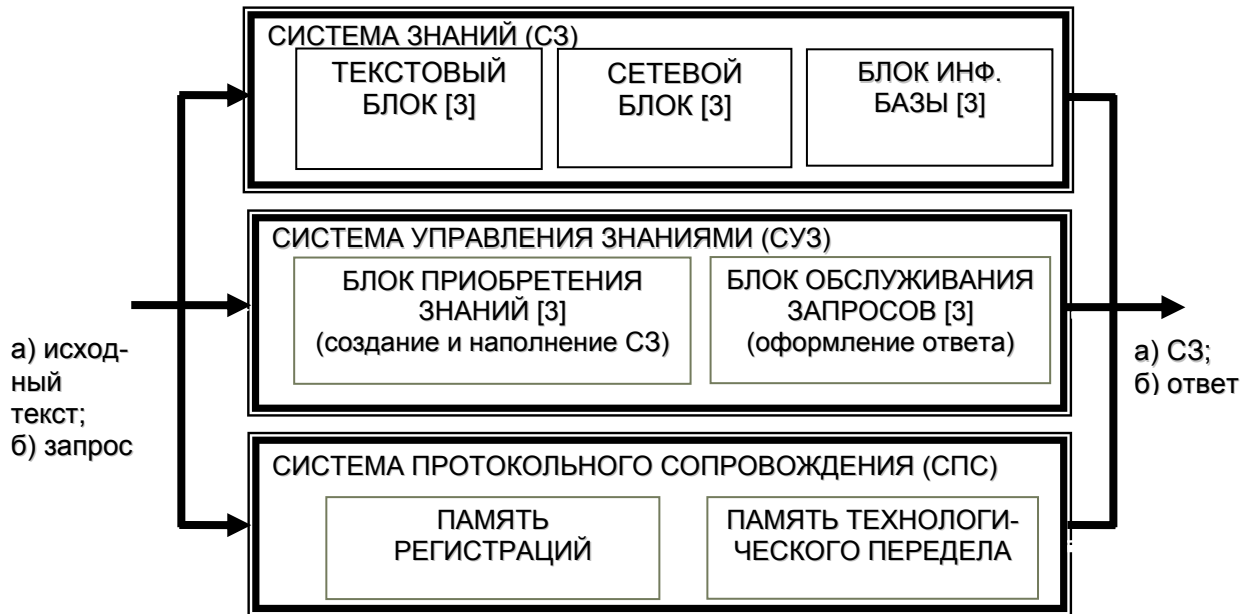


Рис. 1. Укрупненная структурная схема СНОЗ СИП

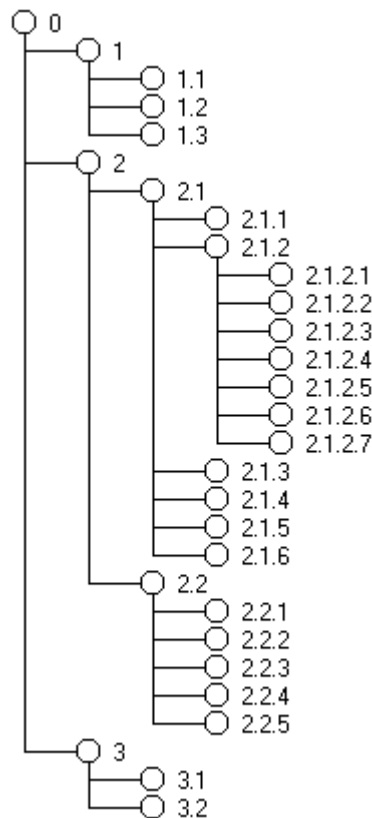


Рис. 10. Иерархическая декомпозиция СНОЗ СИП

(0 – СНОЗ СИП; 1 – СЗ; 1.1 – текстовый блок; 1.2 – сетевой блок; 1.3 – блок информационной базы; 2 – СУЗ; 2.1 – блок приобретения знаний; 2.1.1 –

узел разбиения текстов на предложения [3]; 2.1.2 – узел машинного понимания текстов [3]; 2.1.2.1 – центр указания ключевых терминов и их словооснов [3]; 2.1.2.2 – анализатор вхождения терминов в предложения [2,3]; 2.1.2.3 – центр расчета коэффициентов ассоциативности [2,3]; 2.1.2.4 – центр выбора порога значимости ассоциативной связи [2,3]; 2.1.2.5 – генератор семантических структур [2,3]; 2.1.2.6 – построитель микротезаурусной семантической сети [2,3]; 2.1.2.7 – построитель макротезаурусной семантической сети [2,3]; 2.1.3 – адресатор семантических структур [2,3]; 2.1.4 – сборщик наполняющих текстов [2,3]; 2.1.5 – сборщик информационной базы [2,3]; 2.1.6 – генератор структуры для расширения запросов [2,3]; 2.2 – блок обслуживания запросов; 2.2.1 – узел указания терминов запроса; 2.2.2 – расширитель запроса [2]; 2.2.3 – определитель релевантных семантических структур [2,3]; 2.2.4 – генератор прямого текстового ответа [2,3]; 2.2.5 – определитель релевантной части информационной базы [2,3]; 3 – СПС; 3.1 – память регистраций; 3.2 – память технологического передела)

Задачи, решаемые СНОЗ СИП:

- предоставление пользователю экранных (табличных) форм для переноса в них исходного текста (по предложениям), записи содержащихся ключевых терминов и их словооснов;
- построение матрицы вхождений терминов в предложения;
- расчет коэффициентов ассоциативности [2,3] для терминов;
- выбор порога значимости ассоциативной связи [2,3] (при наличии ручных операций с участием пользователя);
- построение дуплексных (парных) ассоциативных семантических структур [2,3];
- построение семантической сети ассоциирующихся терминов;
- построение онтологической информационной базы [2-6] (создание ее вершин и дуг путем классификации вершин ранее построенной семантической сети по отношению ассоциативности и соответствующей классификации дуг, адресация семантических структур на вершины и дуги созданной информационной базы [2,3], создание текстовых наполнителей вершин и дуг информационной базы на основе адресованных на них семантических структур [2,3]);
- создание вспомогательной структуры для расширения будущих естественно-языковых запросов [2,3] (во избежание излишней сжатости ответов);
- предоставление пользователю экранной (табличной) формы для указания ключевых терминов, одновременно входящих в запрос и отфильтрованных из хранимого текста;
- расширение запроса (т.е. добавление ключевых терминов, отфильтрованных из хранимого текста, не входящих в запрос, но объявленных эквивалентными присутствующим в нем) при необходимости;

- указание терминов, релевантных запросу (входящих в него, либо найденных путем расширения запроса);
- указание семантических структур, релевантных запросу [2,3];
- формирование прямого текстового ответа на запрос [2,3];
- указание вершин и / или дуг информационной базы, релевантных запросу [2,3];
- указание семантических структур, адресованных на каждую из релевантных вершин и / или дуг информационной базы;
- считывание текстовых наполнителей релевантных вершин и / или дуг информационной базы с формированием фокусов внимания (другой шрифт, цвет, фон и т.п.) в каждом из них.

Степень автоматизации операций в созданной исследовательской и перспективной рабочей версии СНОЗ СИП показана в табл.

Таблица

Автоматизация операций в различных версиях СНОЗ СИП

Группы операций	Степень автоматизации операций по версиям	
	Исследовательская	Рабочая
Разбиение текста на предложения	0	0,9
Запись ключевых терминов	0	1
Запись словооснов	0	1
Построение матрицы вхождений терминов в предложения	1	1
Расчет коэффициентов ассоциативности	1	1
Выбор порога значимости ассоциативной связи	0,4	0,8
Построение ассоциативных дуплексных семантических структур	1	1
Построение семантической сети ассоциирующихся терминов	1	1
Создание онтологической информационной базы	0,9	0,9
Просмотр информационной базы	0,6	0,6
Построение структуры для расширения запросов	1	1
Обработка запросов	0,9	0,9

В качестве средств хранения и обработки информации предложено использовать реляционные базы данных в табличном представлении (формат dbf).

Для представления СЗ предложено использовать таблицы: предложений исходного текста; терминов текста и их словооснов; классов ассоциативности и их имен; вершин и дуг информационной базы; семантических структур с адресацией на элементы информационной базы; структуры для расширения запросов. Для обеспечения протокольного сопровождения предложено использовать также таблицы: регистрации пользователей; вхождения терминов в предложения; коэффициентов ассоциативности; выбора порога значимости ассоциатив-

ной связи; исходной (микротезаурусной, неонтологической) семантической сети (в форме матрицы смежности).

Основной риск, сопровождающий дальнейшую разработку СНОЗ СИП, сопряжен по нашему мнению с очень большим объемом памяти (до сотен терабайт), требуемой для хранения и поиска информации о семантических структурах. Мы посчитали возможным решить эту проблему за счет существенного ограничения (до трех страниц) объемов обрабатываемых текстов (что делает возможным ограничиться использованием только таблиц формата dbf), но только в исследовательской версии. Для реализации последующих версий необходим эффективный сжиматель соответствующей информации, например, предложенный В.И. Бодякиным [7], позволяющий хранить заданный текст вместе с семантическими структурами в информационном коде (по типу генетического) и осуществлять поиск необходимых структур путем разворачивания («проигрывания») указанного кода на нейроимитаторе.

СПИСОК ЛИТЕРАТУРЫ:

1. Ткаченко Т.Я. Инструментальная среда системотехнического обслуживания сложных объектов. – Екатеринбург: ГОУ ВПО «УГТУ-УПИ», 2002. – 203с.
2. Гольдштейн С.Л., Кудрявцев А.Г. Разрешение проблемных ситуаций при поддержке систем, основанных на знаниях: Учеб. пособие. – Екатеринбург: ИД «ПироговЪ», 2006. – 218 с.
3. Гольдштейн С.Л., Кудрявцев А.Г. Структура и технологии системного интеллектуального подсказчика по разрешению проблемных ситуаций // Наука и производство: Сборник научных трудов. – Челябинск: ЧНЦ РАН, 2007. – С. 236 -255.
4. Овдей О.М., Проскудина Г.Ю. Обзор инструментов инженерии онтологий; www.rcdl.ru/papers/2005/sek3_2_paper.pdf
5. Гаврилова Т.А., Хорошевский В.Ф. Базы знаний интеллектуальных систем: Учеб. пособие для вузов. - СПб.; М.; Харьков; Минск: ПИТЕР, 2000. - 384 с.
6. Романов А.Н., Одинцов Б.Е. Советующие информационные системы в экономике: Учеб. пособие для студентов вузов. - М.: ЮНИТИ-ДАНА, 2000. - 487 с.
7. Бодякин В.И. Ответы на вопросы совещания; www.ni.iont.ru/N107/WS07/bodyakin.pdf